Background - CGN Deployment

- CGNs may be deployed using a phased hybrid approach
  - Deploy CGNs regionally to start
  - Add CGNs at large headends as needed
  - Rule-of-thumb: 50K subscribers/CGN
CGN Logging Generates Huge Data Volumes

- Subscribers generate around 33,000 connections/day
  - 150-byte log messages x 33,000 connections/day = 5MB/day/sub
  - 1M-sub ISP will generate 150 TB of data/month (1.8 PB/yr)
  - Assuming 50,000 subs/CGN => 23 Mbps of logging traffic
    - PER CGN!!!
  - Imagine DB search times...

- Conclusion: per-connection CGN logging is not scalable
## Ports per subscriber

- **WAND (NZ) Research on CGNs**

<table>
<thead>
<tr>
<th></th>
<th>Outbound</th>
<th>Inbound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Mean Ports/active subscriber</td>
<td>~700 (300 TCP/400 UDP)</td>
<td>~200 (60 TCP/140 UDP)</td>
</tr>
<tr>
<td>Max Median Ports/active subscriber</td>
<td>~70 (40 TCP/10 UDP)</td>
<td>~5</td>
</tr>
<tr>
<td>Peak TCP Ports/active subscriber</td>
<td>99% &lt;1000 90% &lt;100</td>
<td>~15</td>
</tr>
<tr>
<td>Peak UDP Ports/active subscriber</td>
<td>99% &lt;2000 95% &lt;1000 85% &lt;100</td>
<td>~12</td>
</tr>
</tbody>
</table>
Deterministic CGN

• Uses an algorithm to allocate ports per subscriber
  – Abuse Response reverses algorithm to identify inside address
  – Eliminates the need for most logging (except for power users)
  – Reduces privacy impact on subscribers

• Logging only required for subs that exceed predefined port ranges
  – Uses bulk port logging for overflow log
Deterministic CGN Illustrated

DHCP

Reserved Port (e.g. 80)

CGN Device

CGN Mapping Table

IP 1, Port Pool 1
IP 1, Port Pool 2
IP 1, Port Pool 3
IP 1, Port Pool 4
IP 1 Bulk Pool

Static, PCP, portal, etc.

Pool exhausted

Logging Required

Subscriber 1 (DHCP Address 1)
Subscriber 2 (DHCP Address 2)
Subscriber 3 (DHCP Address 3)
Subscriber 4 (DHCP Address 4)
Example

- Assume 100 users; 33k connections/day/user; WAND power curve distribution
- Deterministic – 10:1 compression (~6k ports/subscriber)
- Bulk port assignment – 100 ports/block

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Bulk</th>
<th>Deterministic</th>
<th>SD-NAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>85&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td></td>
<td>85</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>90&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td></td>
<td>25</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td></td>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td></td>
<td>100</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3,300,000</td>
<td>340</td>
<td>40</td>
<td>0</td>
</tr>
</tbody>
</table>

- Bulk - ~10,000x reduction
- Deterministic – ~100,000x reduction
Comparing log reduction approaches

- SD NAT
- Deterministic CGN
- Bulk Port
- Per-connection

Port assignment efficiency vs. Log Volume
Analysis

• bulk port logging, deterministic CGN, and SD-NAT significantly reduce log volumes
  – Bulk log messages approach number of active subs
  – Deterministic log messages approach 0
  – SD-NAT requires no logging

• This draft also compatible with other two approaches
  – Initial ports/user=0: bulk port
  – Overflow buffer=0: SD-NAT
Next Steps

• Update draft with mailing-list feedback
• WG draft?